DOCUMENT RESUME

ED 449 466 CS 014 221

| | |
|---|---|
| TITLE | Comparison of the STAR Reading[R] Computer-Adaptive Test and the Scholastic Reading Inventory[TM]--Interactive Test Report. |
| INSTITUTION | School Renaissance Inst., Inc., Madison, WI. |
| PUB DATE | 2000-05-00 |
| NOTE | 6p. |
| AVAILABLE FROM | School Renaissance Institute, P.O. Box 45016, Madison, WI 53744-5016. Tel: 800-200-4848 (Toll Free); Fax: 608-664-3882; Web site: http://www.SchoolRenaissance.com. |
| PUB TYPE | Reports - Research (143) |
| EDRS PRICE | MF01/PC01 Plus Postage. |
| DESCRIPTORS | *Adaptive Testing; Comparative Analysis; *Comparative Testing; Elementary Secondary Education; Reading Research; *Reading Tests |

ABSTRACT

A study evaluated comparatively the Scholastic Reading Inventory (SRI) Interactive Test and Advantage Learning Systems' STAR Reading Computer-Adaptive Standardized Test. Due to the different methods used for collecting and calculating norm-referenced scores in the two tests, scale score measures of reading performance were used for the comparative study. During January 2000, approximately 10 students per grade, in grades 1-11, were tested once in STAR Reading version 2.0 (STAR) and twice in SRI-Interactive version 1.1 (SRI). Children were tested twice; during the first test session, children took a STAR test first and an SRI test second, while during the second session, children took only a second SRI test. Results showed that the SRI test takes twice as long and is significantly less reliable than STAR. The conclusion is that the STAR test is superior in all measurable respects important to teachers and students. (Contains 6 tables of data.) (NKA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 449 466

# Comparison of the STAR Reading Computer-Adaptive Test and the Scholastic Reading Inventory-Interactive Test.

Report

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

K. Hintz

1



CS 014 221

REPORT from the School Renaissance Institute

May 2000

# Comparison of the STAR Reading® Computer-Adaptive Test and the Scholastic Reading Inventory™–Interactive Test

The following study presents the results of a comparative evaluation of the Scholastic Reading Inventory (SRI)–Interactive test and Advantage Learning Systems' STAR Reading® computer-adaptive standardized test.

## Background

STAR Reading version 2.0 is a fixed-length, 25-item, computer-adaptive test. For most students, the test consists of both vocabulary-in-context items and authentic text passage items. Using standard test development methods, STAR Reading 2.0 was normed in the spring of 1999 by administering the test to a nationally representative sample of 30,000 students in 269 schools in 47 states. STAR Reading uses the data collected from the norming sample directly to report norm-referenced scores, such as percentile rank (PR) and normal curve equivalent (NCE).

SRI–Interactive version 1.1 is a variable-length test. It presents extended passage items to students of all grade levels. While it provides PR, NCE, and other norm-referenced scores, the test itself was not nationally normed. Instead, norm-referenced scores are based on a scale linkage between SRI–Interactive and some other standardized test. Data for the scale linkage were collected from a limited, non-representative sample in one state.

Due to the different methods used for collecting and calculating norm-referenced scores in the two tests, scale score measures of reading performance were used for this comparative study.

## The Testing Procedure

During January 2000, approximately ten students per grade, in grades one through eleven, were tested once in STAR Reading version 2.0 (STAR) and twice in SRI–Interactive version 1.1 (SRI). Parents signed their child(ren) up for two test sessions, each on a different day, but separated by no more than two weeks. During the first test session, children took a STAR test first and an SRI test second. During the second test session, children took only a second SRI test.

Each child was assigned to a different test station for each test session in order to prevent retest effects within the software. A test monitor was assigned to each test station and each child, and was responsible for reading the pretest instructions, monitoring the test time, and recording observations. The test monitor also administered a short survey to the child at the end of the first test session.

Of the 109 students originally in the study, four were eliminated because they had extreme difficulty in reading the SRI test questions on their own (three first graders and one second grader) and one was deleted because he didn't come back for a second SRI test (a fifth grader). Eliminating these five left 104 valid students, each with one STAR test and two SRI tests.

The data were analyzed to meet the following objectives:

- Compare the average overall test time of STAR and SRI tests.
- Determine the amount of variance in overall test time of the two tests.
- Compare the test score distributions of the STAR and SRI tests.
- Measure the test-retest reliability of the SRI test and compare it to that of STAR.

## Overall Test Time

The STAR test took 8.3 ± 2.3 minutes, for an average range of 6.0 to 10.6 minutes.

By comparison, the first SRI test in this study took 17.9 ± 6.4 minutes (11.5 to 24.3 minutes), and the second SRI test took 16.5 ± 5.5 minutes (11.0 to 22.0 minutes). SRI tests therefore took about twice as long as STAR tests and the test time was about three times more variable than STAR. This time variability is likely due to SRI being a variable-length test consisting of only extended passage items.

## Test Score Distributions

Summary statistics for individual test scores—Scaled Score for STAR and Lexile® Score for the two SRI tests—by test and by grade, are given in Tables 1a, 1b, and 1c.

**Table 1a: STAR Test Score (Scaled Score) Distribution**

| Grade | N | Mean | S.D. | Min. | Max. |
|---|---|---|---|---|---|
| 1 | 4 | 340 | 118 | 259 | 509 |
| 2 | 10 | 429 | 129 | 250 | 591 |
| 3 | 11 | 514 | 97 | 395 | 758 |
| 4 | 9 | 673 | 118 | 496 | 813 |
| 5 | 13 | 808 | 230 | 541 | 1259 |
| 6 | 10 | 843 | 225 | 570 | 1163 |
| 7 | 12 | 992 | 258 | 644 | 1347 |
| 8 | 9 | 1164 | 131 | 855 | 1308 |
| 9 | 11 | 1150 | 189 | 803 | 1347 |
| 10 | 6 | 1220 | 191 | 906 | 1345 |
| 11 | 9 | 1161 | 298 | 607 | 1347 |
| **All** | **104** | **857** | **338** | **250** | **1347** |

**Table 1b: SRI #1 Test Score (Lexile Score) Distribution**

| Grade | N | Mean | S.D. | Min. | Max. |
|---|---|---|---|---|---|
| 1 | 4 | 387 | 214 | 237 | 699 |
| 2 | 10 | 457 | 298 | 100 | 814 |
| 3 | 11 | 652 | 206 | 343 | 949 |
| 4 | 9 | 897 | 190 | 585 | 1241 |
| 5 | 13 | 992 | 155 | 739 | 1306 |
| 6 | 10 | 1000 | 182 | 705 | 1357 |
| 7 | 12 | 1124 | 214 | 706 | 1487 |
| 8 | 9 | 1220 | 177 | 920 | 1490 |
| 9 | 11 | 1209 | 221 | 801 | 1500 |
| 10 | 6 | 1239 | 210 | 955 | 1500 |
| 11 | 9 | 1271 | 187 | 980 | 1500 |
| **All** | **104** | **970** | **340** | **100** | **1500** |

**Table 1c: SRI #2 Test Score (Lexile Score) Distribution**

| Grade | N | Mean | S.D. | Min. | Max. |
|---|---|---|---|---|---|
| 1 | 4 | 385 | 197 | 100 | 551 |
| 2 | 10 | 461 | 294 | 100 | 842 |
| 3 | 11 | 605 | 190 | 246 | 869 |
| 4 | 9 | 859 | 126 | 720 | 1046 |
| 5 | 13 | 952 | 193 | 561 | 1224 |
| 6 | 10 | 973 | 186 | 543 | 1254 |
| 7 | 12 | 1101 | 201 | 717 | 1377 |
| 8 | 9 | 1222 | 294 | 910 | 1500 |
| 9 | 11 | 1220 | 166 | 977 | 1500 |
| 10 | 6 | 1285 | 188 | 1083 | 1500 |
| 11 | 9 | 1245 | 296 | 557 | 1489 |
| **All** | **104** | **954** | **346** | **100** | **1500** |

As shown in Tables 1b and 1c, students frequently received Lexile Scores in the 800 to 1500 range, at the upper end of the 100 to 1500 Lexile scale. The maximum Lexile Score for eighth- through eleventh-graders was, in fact, 1500. In contrast, students' Scaled Scores on the STAR test (Table 1a) were evenly distributed over the 0 to 1400 Scaled Score range. This even distribution of scores indicates that STAR is a more accurate measure of reading ability at all grade and ability levels.

## SRI Test-Retest Reliability

The correlation, R, between students' Lexile Scores on the two SRI tests is given in Table 2.

**Table 2: SRI Test-Retest Correlation Coefficients**

| N | R | $R^2$ | S.E. |
|---|---|---|---|
| 104 | 0.886 | 0.784 | 158.4 |

The test-retest reliability of SRI tests in this sample was therefore approximately 0.89. This is significantly less than the estimated test-retest reliability calculated for STAR in the 1999 norming sample (N=2,095), 0.94.

## Observation Data

Test monitors observed students during all three tests and took notes about each student's attitude, level of engagement, comments, and relative ease or difficulty in taking each test. Observations were recorded during 28 of the STAR tests (27% of cases), 27 of the first SRI tests (26% of cases), and 18 of the second SRI tests (17% of cases). Table 3 lists the number and percentage of observations that fell into two main categories.

**Table 3: Main Observation Categories**

| The student ... | STAR | SRI #1 | SRI #2 |
|---|---|---|---|
| Seemed calm, at ease, relaxed, comfortable, confident | 9 (32%) | 2 (7%) | 1 (6%) |
| Seemed bored, distracted, tired, uninterested, fidgety | 4 (14%) | 14 (52%) | 5 (28%) |

In 32% of the observations recorded during the STAR test, it was noted that the student seemed calm, at ease, relaxed, comfortable, or confident. Such positive student reactions were only noted in 7% of the observations recorded during the first SRI test. Likewise, in 52% of the observations recorded during the first SRI test, it was noted that the student seemed bored, distracted, tired, uninterested, or fidgety. Negative student reactions such as these were only noted in 14% of the observations recorded during the STAR test.

## Survey Data

At the end of the first session, which involved taking a STAR test first and an SRI test second, students were asked the following multiple-choice question:

*Which of the following statements do you agree with?*

(1) The *first* test was **much easier** than the second test.

(2) The *first* test was **a little easier** than the second test.

(3) The *first* test was **about the same** as the second test.

(4) The *second* test was **a little easier** than the first test.

(5) The *second* test was **much easier** than the first test.

The percentages of students who chose each answer are shown below in Table 4.

**Table 4: Percentage of Students Who Chose Each Answer**

| N | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 101 | 43% | 26% | 16% | 11% | 4% |

The majority of students (69%) thought that the STAR test was much easier or a little easier than the SRI. Only 15% thought that the SRI test was easier than STAR.

Students were also asked the following two open-ended questions:

*What did you like or dislike about the first test?*

*What did you like or dislike about the second test?*

Eighty-eight students responded to the first open-ended question, and 94 students responded to the second.

Twenty-eight percent of respondents liked the shorter items on the STAR test and disliked the longer items on the SRI test. Several students mentioned that they liked the shorter, fixed-length STAR test. On the SRI test, they liked the ability to skip up to three questions, but they tended not to use their "skip allowance" because, on any given item, they didn't know how many more items remained.

Many students noted that the words in SRI answer choices seemed to all have similar meanings. They thought this made it hard to concentrate and choose the right answer on the SRI test.

Several students liked the story content or format of the SRI test. Before starting the actual SRI test, the SRI software asked each student to pick the types of books he or she liked to read. Some students thought that only questions from the categories they picked would be on the test. They were disappointed when this was not the case.

## Conclusions

The results of this evaluation show that the SRI test takes twice as long and is significantly less reliable than STAR. Students frequently obtained Lexile Scores in the maximum range of 800 to 1500—therefore, the test did not appear to have sufficient "top" to it for discerning the reading levels of upper-grade or higher-ability students. Students in the lower grades or of lower reading ability often had extreme difficulty answering SRI test questions and they quickly became distracted or bored. Many students disliked the SRI test for the length of its reading passages, the similarity in meaning in its answer choice words, and the seemingly unending nature of the test as a whole.

In conclusion, the STAR test is superior in all measurable respects important to teachers and students.

For more information, or for additional copies of this report, write or call:

School Renaissance Institute
P.O. Box 45016
Madison, WI 53744-5016
(800) 200-4848

---

STAR Reading is a registered trademark of Advantage Learning Systems, Inc. Lexile is a registered trademark of MetaMetrics, Inc.
© 2000, The School Renaissance Institute, Inc.
All rights reserved.

L0361.500.ND.20M